

## Penerapan *algorithm* *k-mean* untuk *clustering* data obat pada puskesmas rumbai

Taslim<sup>1</sup>, Fajrizal<sup>2</sup>

<sup>12</sup>Program Studi Teknik Informatika, Universitas Lancang Kuning, Pekanbaru

Jl. Yos Sudarso KM 8 Rumbai, Pekanbaru

E-mail: <sup>1</sup>taslim.malano@gmail.com, <sup>2</sup>Fajrizal@unilak.ac.id

### **Abstrak**

*Melalui program jaminan kesehatan pemerintah berupaya terus menjamin kesehatan bagi masyarakat melalui puskesmas puskesmas atau balai pengobatan. Salah satu komponen yang sangat penting pada puskesmas maupun balai pengobatan adalah masalah ketersediaan obat. Ketersediaan obat harus dikelola secara baik untuk menjamin obat yang dibutuhkan oleh masyarakat selalu tersedia dengan jumlah yang cukup dan memadai. Clusterisasi pada data mining dapat digunakan untuk menganalisa pemakaian obat yang terjadi selama ini pada sebuah puskesmas untuk digunakan sebagai salah satu alat bantu penunjang keputusan bagi pihak puskesmas untuk mengajukan permintaan obat pada periode yang akan datang. Hasil dari penelitian ini dapat mengelompokkan tingkat pemakaian obat pada apotik puskesmas Rumbai Bukit Pekanbaru*

**Kata Kunci :** *Data mining, k-means, obat, puskesmas, cluster*

### **Abstrack**

*Through the government health insurance program seeks to continue to ensure public health through public health centers health centers or clinics. One very important component in health centers and clinics is the issue of availability of drugs. Availability of the drug should be managed properly to ensure the drug is needed by the community always available in sufficient quantities and adequate. Clusterisasi on data mining can be used to analyze the use of drugs that occurred during this time in a health clinic to be used as a decision support tool for the clinic to request the drug in the coming period. The results of this research can classify the level of drug use at the health center pharmacy Tassel Hill Pekanbaru*

**Keyword:** *Data mining, k-means, medicine, health centers, cluster*

### **1. Pendahuluan**

Obat merupakan salah satu komponen yang penting dalam hal kesehatan baik untuk mencegah, mengurangi, menghilangkan atau menyembuhkan suatu penyakit atau gejala penyakit. Untuk itu obat perlu dikelola dengan baik, efektif dan efisien. Perencanaan kebutuhan obat merupakan hal yang penting dilakukan untuk menjamin ketersediann dan pemerataan obat dengan jenis dan jumlah yang mencukupi agar obat dapat diperoleh dengan cepat pada tempat dan waktu yang tepat pada instansi instansi yang terkait dengan pelayanan kesehatan, baik itu rumah sakit, puskesmas, dinas kesehatan dan lain sebagainya. Perencanaan kebutuhan obat akan mempengaruhi pengadaan, pendistribusian dan dan pemakaian obat di tempat pelayanan kesehatan

*Clusterisasi* kebutuhan obat diharapkan dapat menjadi salah satu sumber pengambilan keputusan untuk menjamin ketersediaan obat pada dinas kesehatan Kota Pekanbaru. Clustering data merupakan salah satu metoda dalam data mining yang dapat digunakan untuk memetakan data kedalam kelompok kelompok yang lebih kecil berdasarkan kesamaan karakteristik yang dimilikinya (Perim, Wandekokem, & Varejão, 2008). Dengan hasil *clustering* ini distribusi obat

---

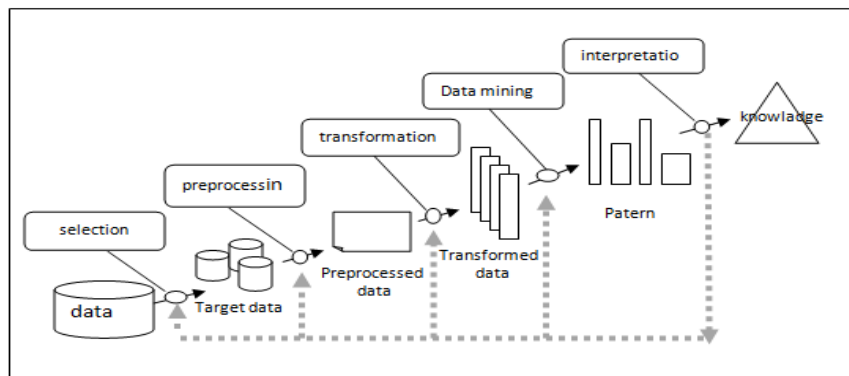
pada instansi layanan kesehatan dapat dikelompokkan sesuai kebutuhan berdasarkan data distribusi obat pada tahun tahun sebelumnya dan dapat digunakan sebagai acuan perencanaan obat untuk tahun berikutnya. Dengan begitu diharapkan ketersediaan obat untuk tahun berikutnya dapat lebih terjamin dan dapat memenuhi permintaan permintaan obat dari instansi instansi layanan kesehatan.

Salah satu metoda clustering yang paling terkenal diantara algoritma clustering yaitu K-means (Patel & Mehta, 2011). Kesederhanaan metode ini membuat algoritma *K-means* dapat diaplikasikan pada berbagai bidang (K.Arai and A.R. Barakbah, 2007).

## 2. Metode Penelitian

### 2.1 Tahapan tahapan dalam data mining

KDD adalah proses nontrivial mengidentifikasi validitas data, potensi, guna, dan akhirnya menghasilkan pola data yang dapat dimengerti (Daniel t. larose 2007). Adapun tahapan tahapan dalam data mining yaitu :



Gambar 1. Tahapan tahapan pada data mining

Pada penelitian ini digunakan data yang berasal dari Laporan Pemakaian dan Lembar Permintaan Obat (LPLPO) Puskesmas Rumbai Bukit tahun 2014. Selanjutnya dilakukan proses *clustering* terhadap data yang ada untuk mendapatkan pola kebutuhan obat untuk masyarakat wilayah Rumbai Pekanbaru.

### 2.2 Clustering

Clustering mengacu pada pengelompokan catatan, observasi, atau kasus ke dalam kelas yang serupa. Sebuah cluster adalah kumpulan dari catatan yang mirip satu sama lain dan berbeda dengan catatan dalam cluster lain. *Clustering* berbeda dengan *classification* yang tidak ada variabel target untuk clustering. Sebaliknya, algoritma klasterisasi mencari ke segmen data seluruh set menjadi sub kelompok yang relatif homogen atau kelompok, di mana kesamaan catatan dalam cluster dimaksimalkan, dan kesamaan catatan luar klaster ini diminimalkan (Daniel t. Larose, 2005).

Contoh tugas pengelompokan dalam bisnis dan penelitian meliputi :

1. Menentukan target pemasaran produk
2. Untuk tujuan audit akuntansi, melihat perilaku keuangan yang mencurigakan
3. Sebagai alat dimensi pengurangan ketika data set memiliki ratusan atribut
4. Untuk pengelompokan ekspresi gen, di mana jumlah yang sangat besar mungkin menunjukkan perilaku yang sama.

Pada analisis cluster ada beberapa hal yang harus diperhatikan, sebagai contoh, kita akan perlu menentukan:

1. Bagaimana mengukur kesamaan
2. Bagaimana recode variabel kategori

3. Bagaimana standarisasi atau menormalkan variabel numerik
4. Berapa banyak cluster yang akan dibentuk

### 2.3 Interval-skala variabel

Interval-skala variabel adalah pengukuran kontinyu yang meliputi skala linier. Contoh, berat dan tinggi, lintang dan bujur dan suhu cuaca. Penggunaan unit pengukuran dapat mempengaruhi analisa cluster[8]. Sebagai contoh perubahan pengukuran dari meter ke inchi, atau dari kilogram ke pound, mungkin akan menghasilkan stuktur cluster yang sangat berbeda. Untuk itu diperlukan normalisasi data, agar semua data mempunyai bobot yang sama. Langkah langkah untuk normalisasi data yaitu :

1. Hitung nilai mean

$$sf = \frac{1}{n} ([x1f - mf] + [x2f - mf] + \dots + [xnf - mf])$$

2. Hitung nilai z-score

$$zif = \frac{xif - mf}{sf}$$

### 2.4 Algoritma K-Means

Algoritma k -means Clustering[1] adalah algoritma sederhana dan efektif untuk menemukan cluster dalam data dengan algoritma sebagai berikut:

- Langkah 1 : Tentukan jumlah cluster
- Langkah 2 : Tentukan nilai yang menjadi lokasi pusat cluster awal .
- Langkah 3 : Hitung pusat cluster terdekat untuk setiap record
- Langkah 4 : Untuk setiap cluster k , hitung centroid cluster dan memperbarui lokasi setiap pusat cluster
- Langkah 5 : Ulangi langkah 3 sampai 5 sampai konvergensi atau penghentian .

Algoritma k-means dikenal dan banyak digunakan untuk metoda partitional, yaitu membagi himpunan objek data ke dalam sub himpunan cluster yang tidak saling tumpang tindih, sehingga setiap objek data tepat berada dalam satu *cluster*

## 3. Hasil dan Pembahasan

Data sumber yang digunakan pada penelitian ini adalah data yang berasal LPLPO Puskesmas Rumbai Bukit tahun 2014. Adapun data yang digunakan dapat dilihat pada tabel 1.

### REKAPITULASI RESEP APOTEK

PUSKESMAS : RUMBAI BUKIT  
BULAN : JANUARI  
TAHUN : 2014

NO	NAMA OBAT	Satuan	Bulan											
			Jan	Feb	Maret	April	Mei	Juni	Juli	Agust	Sept	Okt	Nov	Des
1	Acyclovir krim	tube	4	5	4	21	23	15	19	19	24	24	91	75
2	Acyclovir 200 mg	tablet	20	110	20	95	110	132	23	0	0	0	0	0
3	Acyclovir 400 mg	tablet	10	0	0	488	0	0	210	100	200	195	220	115
4	Alat suntik 1 ml	set	0	0	0	0	0	0	0	0	100	0	0	0
5	Alat suntik 5 ml	set	0	0	0	0	0	0	0	0	0	0	100	0
6	Albendazol 400 mg tab	tablet	23	54	13	6	32	87	15	15	16	6	66	60
7	Alluporinol tab 100 mg	tablet	20	170	150	60	70	270	50	493	0	0	0	0
8	Ambroxol 30 mg	tablet	191	200	3028	225	127	72	120	20	149	340	1602	1035
9	Ambroxol sirup	botol	72	56	57	45	28	22	50	30	193	0	0	0
10	Aminofilin tab 200 mg	tablet	10	0	152	0	72	10	0	0	40	0	45	33

Terdapat 133 jenis obat pada apotik puskesmas rumbai bukit, adapun sebagian dari jenis obat dan pengeluaran obat tiap bulan dapat dilihat pada table berikut

Tabel 2. Rekapitulasi Obat Perbulan

NO	NAMA OBAT	Satuan	Bulan											
			Jan	Feb	Maret	April	Mei	Juni	Juli	Agust	Sept	Okt	Nov	Des
1	Acyclovir krim	tube	4	5	4	21	23	15	19	19	24	24	91	75
2	Acyclovir 200 mg	tablet	20	110	20	95	110	132	23	0	0	0	0	0
3	Acyclovir 400 mg	tablet	10	0	0	488	0	0	210	100	200	195	220	115
4	Alat suntik 1 ml	set	0	0	0	0	0	0	0	0	100	0	0	0
5	Alat suntik 5 ml	set	0	0	0	0	0	0	0	0	0	0	100	0
6	Albendazol 400 mg tab	tablet	23	54	13	6	32	87	15	15	16	6	66	60
7	Alluporinol tab 100 mg	tablet	20	170	150	60	70	270	50	493	0	0	0	0
8	Ambroxol 30 mg	tablet	191	200	3028	225	127	72	120	20	149	340	1602	1035
9	Ambroxol sirup	botol	72	56	57	45	28	22	50	30	193	0	0	0
10	Aminofilin tab 200 mg	tablet	10	0	152	0	72	10	0	0	40	0	45	33
11	Amoksisilin tab 250 mg	kapsul	1020	924	844	660	750	1456	1344	780	1142	1308	2002	1474
12	Amoksisilin cap 500 mg	kapsul	2245	2310	2370	1780	1503	1765	1795	1210	4340	2110	2754	2042
13	Amoksisilin syr kering 125 mg	botol	110	110	77	83	96	166	162	91	145	161	193	150
14	Amlodipin besilat 10 mg	tablet	150	130	180	550	755	200	20	40	371	90	400	110
15	Antasida tab doen	tablet	912	2147	1023	1090	1098	1145	990	268	529	693	2030	1724
16	Antasida doen suspensi	botol	37	29	22	37	18	13	49	44	100	62	60	45
17	Anti bakteri doen salap	tube	14	21	20	15	10	19	24	7	12	17	35	25
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
133	Loperamid		94	122	128	20	0	0	0	0	0	0	0	0

Dari data diatas (tabel 2) selanjutnya dilakukan proses normalisasi data agar data mempunyai rentang nilai yang lebih kecil sehingga dapat mempercepat dalam proses perhitungan normalisasi, pada penelitian ini digunakan normalisasi z-score. Sebagian hasil normalisasi dapat dilihat pada table 3 .

Tabel 3. Hasil normalisasi dengan z-score

0.0000 0.0000 0.0004 0.0021 0.0023 0.0015 0.0019 0.0019 0.0024 0.0024  
 0.0091 0.0075  
 0.0000 0.0000 0.0020 0.0095 0.0110 0.0132 0.0023 0 0 0 0  
 0

```

0.0000    0    0 0.0488    0    0 0.0210 0.0100 0.0200 0.0195 0.0220
0.0115
    0    0    0    0    0    0    0    0.0100    0    0    0
    0    0    0    0    0    0    0    0    0    0.0100    0
0.0000 0.0000 0.0013 0.0006 0.0032 0.0087 0.0015 0.0015 0.0016 0.0006
0.0066 0.0060
0.0000 0.0000 0.0150 0.0060 0.0070 0.0270 0.0050 0.0493    0    0    0
0
0.0000 0.0000 0.3028 0.0225 0.0127 0.0072 0.0120 0.0020 0.0149 0.0340
0.1602 0.1035
0.0000 0.0000 0.0057 0.0045 0.0028 0.0022 0.0050 0.0030 0.0193    0
0    0
0.0000    0 0.0152    0 0.0072 0.0010    0    0 0.0040    0 0.0045
0.0033
0.0003 0.0002 0.0844 0.0660 0.0750 0.1456 0.1344 0.0780 0.1142 0.1308
0.2002 0.1474
0.0006 0.0005 0.2370 0.1780 0.1503 0.1765 0.1795 0.1210 0.4340 0.2110
0.2754 0.2042
0.0000 0.0000 0.0077 0.0083 0.0096 0.0166 0.0162 0.0091 0.0145 0.0161
0.0193 0.0150
0.0000 0.0000 0.0180 0.0550 0.0755 0.0200 0.0020 0.0040 0.0371 0.0090
0.0400 0.0110
0.0002 0.0005 0.1023 0.1090 0.1098 0.1145 0.0990 0.0268 0.0529 0.0693
0.2030 0.1724
0.0000 0.0000 0.0022 0.0037 0.0018 0.0013 0.0049 0.0044 0.0100 0.0062
0.0060 0.0045
0.0000 0.0000 0.0020 0.0015 0.0010 0.0019 0.0024 0.0007 0.0012 0.0017
0.0035 0.0025
0.0000 0.0000 0.0014 0.0014 0.0021 0.0014    0    0    0    0    0
0
.....
.....
0.0000 0.0000 0.0128 0.0020    0    0    0    0    0    0    0    0

```

Tahap berikutnya ditentukan jumlah cluster yakni sebanyak 3 cluster, untuk dapat menentukan apakah suatu permintaan obat mempunyai permintaan “sangat tinggi”, “tinggi”, atau “kurang”. Untuk nilai centroid awal dipilih secara acak. Nilai centroid diambil dari baris 50,60 dan 70.

#### Iterasi 1

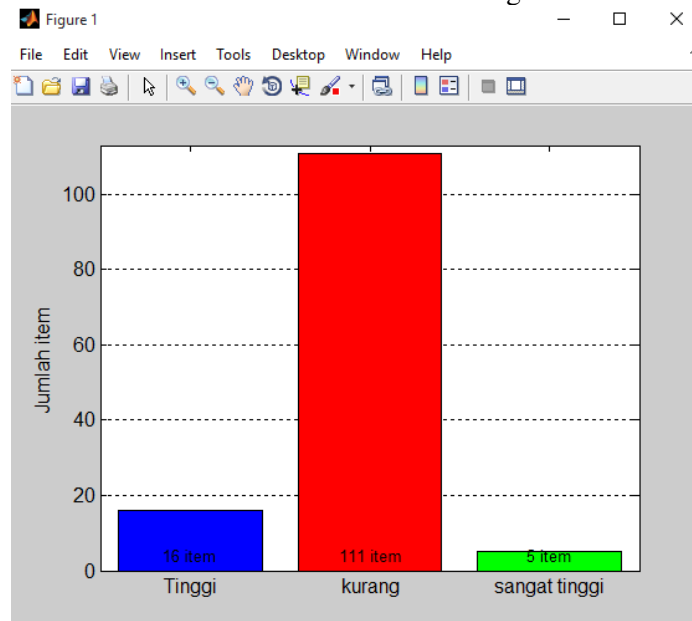
Pada langkah ini, pembaruan pusat pusat kelompok akan dilakukan sebagai berikut (tabel 3).

Tabel 4. nilai centroid awal

0.0157	0.0045	2.0000	1.0000	1.0000	2.0000	1.0000	1.0000	2.0000	2.0000	0	13.0000
0	0	0	0	0	0	0	0	0	0	0	200.0000
0	0	0	10.0000	0	0	0	0	0	230.0000	145.0000	0

Selanjutnya dilakukan proses clustering d0 (tabel 4). Dari hasil clustering setelah 3 kali iterasi didapat hasil pengelompokkan data obat, yang dikelompokkan jadi 3 kelompok yaitu kelompok permintaan obat sangat tinggi, kelompok obat dengan permintaan tinggi dan kelompok permintaan obat yang tergolong kurang. Adapun jumlah hasil clustering dapat dilihat pada table berikut.

Gambar 4. Clustering 1



Sedangkan sebagian rincian table *clustering* data obat dapat dilihat pada table berikut.

Tabel 4. Hasil Clustering Permintaan obat

1	Acyclovir krim	tube	4	5	4	21	23	15	19	19	24	24	91	75
2	Acyclovir 200 mg	tablet	20	110	20	95	110	132	23	0	0	0	0	0
3	Acyclovir 400 mg	tablet	10	0	0	488	0	0	210	100	200	195	220	115
4	Alat suntik 1 ml	set	0	0	0	0	0	0	0	0	100	0	0	0
5	Alat suntik 5 ml	set	0	0	0	0	0	0	0	0	0	0	100	0
6	Albendazol 400 mg tab	tablet	23	54	13	6	32	87	15	15	16	6	66	60
7	Alluporinol tab 100 mg	tablet	20	170	150	60	70	270	50	493	0	0	0	0
8	Ambroxol 30 mg	tablet	191	200	3028	225	127	72	120	20	149	340	1602	1035
9	Ambroxol sirup	botol	72	56	57	45	28	22	50	30	193	0	0	0
10	Aminofilin tab 200 mg	tablet	10	0	152	0	72	10	0	0	40	0	45	33
11	Amoksisilin tab 250 mg	kapsul	1020	924	844	660	750	1456	1344	780	1142	1308	2002	1474
12	Amoksisilin cap 500 mg	kapsul	2245	2310	2370	1780	1503	1765	1795	1210	4340	2110	2754	2042
13	Amoksisilin syr kering 125 mg	botol	110	110	77	83	96	166	162	91	145	161	193	150
14	Amlodipin besilat 10 mg	tablet	150	130	180	550	755	200	20	40	371	90	400	110
15	Antasida tab doen	tablet	912	2147	1023	1090	1098	1145	990	268	529	693	2030	1724
16	Antasida doen suspensi	botol	37	29	22	37	18	13	49	44	100	62	60	45

17	Anti bakteri doen salap	tube	14	21	20	15	10	19	24	7	12	17	35	25
18	Anti fungi doen salep	pot	15	44	14	14	21	14	0	0	0	0	0	0
19	Anti hemoroid supp komb	supp	0	0	0	0	0	13	16	0	3	0	8	0
20	Asam askorbat tab 50 mg	tablet	1419	3812	3783	3005	2784	3392	3890	2703	3955	3567	7090	6112
21	Asam mefenamat tab 500 mg	kaplet	800	940	1040	1032	930	1000	570	420	820	715	1708	1510

#### 4. Kesimpulan

Dari hasil clusterisasi data obat dapat disimpulkan bahwa banyak diantara obat yang ada masuk dalam kelompok cluster kurang. Kecendrungan obat yang tergolong kurang ini adalah ada beberapa bulan yang tidak ada permintaan sama sekali terhadap obat tersebut. Sedangkan kelompok obat yang termasuk tinggi rata rata permintaan obatnya setiap bulan diatas 300 buah, sedangkan obat yang masuk dalam kelompok cluster sangat tinggi rata rata permintaan setiap bulannya adalah diatas 2000 buah. Dari hasil analisa cluster diatas mungkin perlu dilakukan lagi penelitian lanjutan agar *clusterisasi* data obat dapat dilakukan secara lebih valid dengan menetapkan nilai centroid terbaik.

#### Daftar Pustaka

- [1] Tipawan.S and Kulthida.T “Data mining and its application for knowledge management : a literature review from 2007 to 2012 ” (2012)
- [2] Hans-Peter .K, Karsten, Peer.K, Alexey.P, Matthias.S, Arthur.Z, “Future trends in data mining” (2007)
- [3] J.Ranjan “Application of data mining techniques in pharmaceutical industry” (2007)
- [4] K.Arai and A.R. Barakbah “Hierarchical K-means : and algorithm for centroids initialization for K-means”.(2007)
- [5] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth “From Data Mining to Knowledge Discovery in Databases” (1996)
- [6] Daniel t. larose “Discovering knowledge in data an Introduction to Data Mining”, John Wiley & Sons 2005